Earth Observation data and knowledge preservation at ESA

from bit archiving to Earth Science long term applications

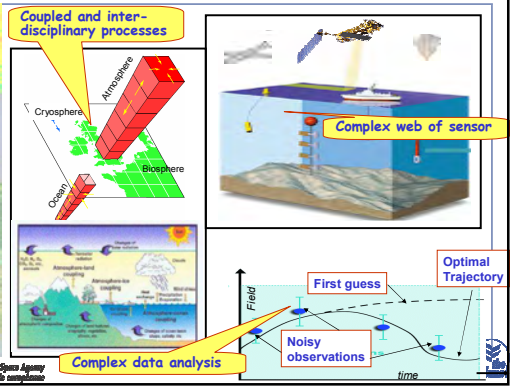
Vinzo Beruti & Luigi Fusco
ESA - Earth Observation Programmes

e-IRG - User needs in IT and Information Infrastructure
Nottingham, 6 Dec 05

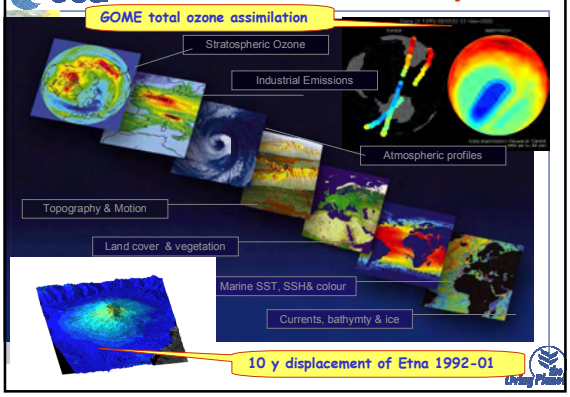
Summary

- ✓The Earth Science world and Earth Observation capabilities
- ✓EO operational scenarios
- ✓EO archiving, preservation and exploitation issues

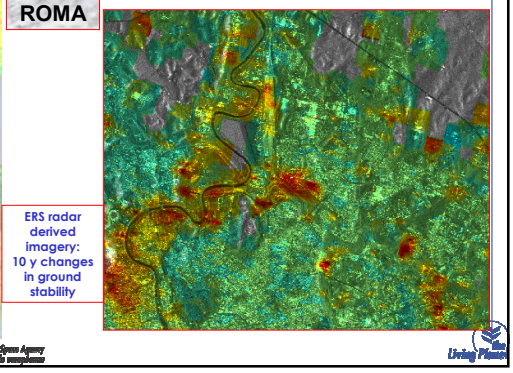
The Earth Science world



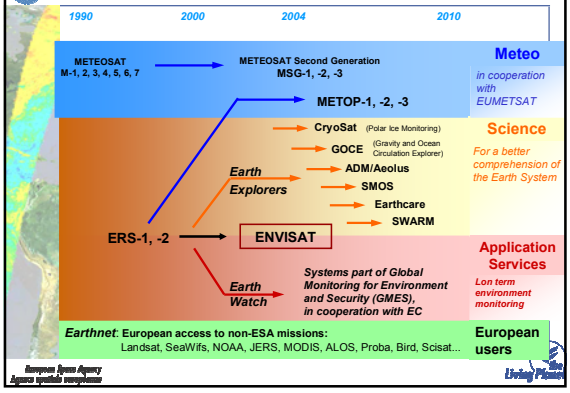
Earth Observation potential



More Earth Observation potential

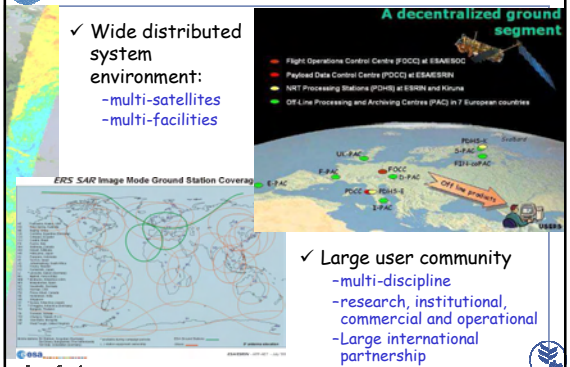


European EO space capacity



The Earth Observation operation scenario

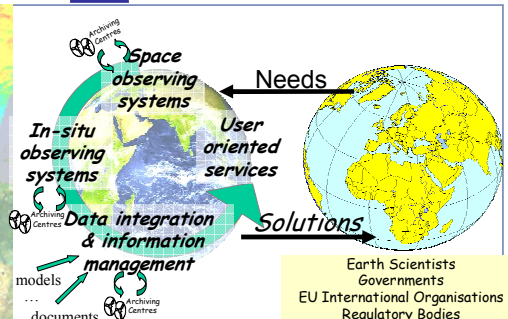
- ✓ Wide distributed system environment:
  - -multi-satellites
  - -multi-facilities
- ✓ Large user community
  - -multi-discipline
  - -research, institutional, commercial and operational
  - -Large international partnership



Some Earth Science digital data requirements

- ✓ Global, regional, local applications
  - – Alternative use of the data at different resolution
- ✓ Large historical distributed archives
  - – Long term data and knowledge preservation issues
- ✓ Near real-time access to data
  - – For processing, value adding and dissemination
- ✓ Integration with models to provide long term trends and forecast
  - – Data assimilation
- ✓ Integrate different data sources
  - – Standardisation, Virtual Organisation, …
- ✓ Need to link data to technical information and scientific results
  - – Need to keep/transfer the historical Knowledge

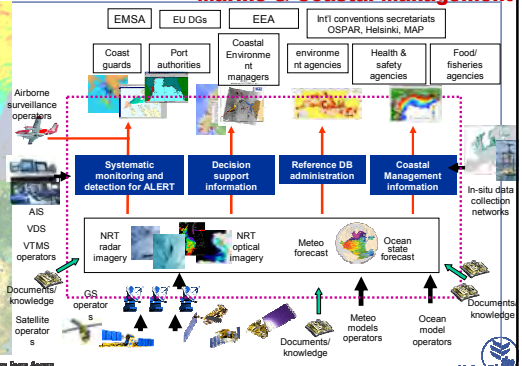
The near future perspective: GMES



Earth Scientists
Governments
EU International Organisations
Regulatory Bodies
Industry
General Public

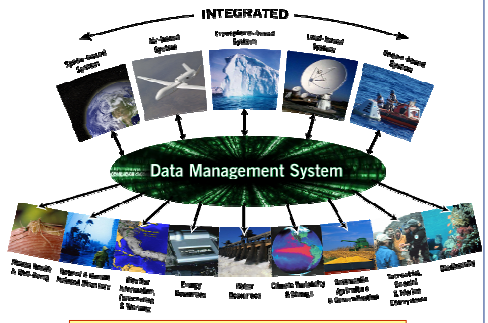
http://www.gmes.info

A GMES example: Marine & Coastal Management



GEOSS – the wider view

Global Earth Observation System of Systems

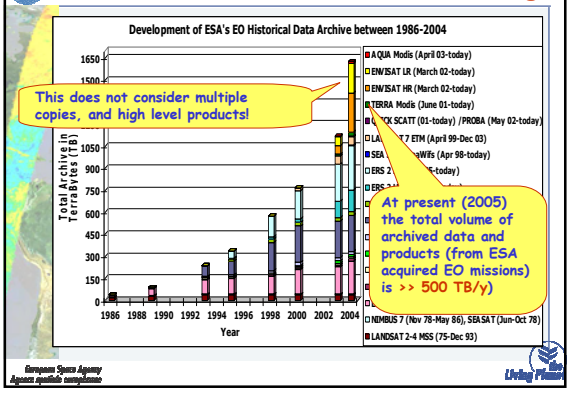


The interoperability framework

The EO infrastructure issues

- ✓ Bring together in the virtual organisation
  - – the users: scientists, institutions …
  - – the space data providers: acquisition stations, processing and archiving facilities
  - – the other data providers
- ✓ Access services … not simple data
  - – Interoperability issues
- ✓ Need for linking all available expertise and knowledge in data handling/usage

EO data volumes are increasing...



Large no of facilities involved...

- Historical missions
  - SPOT 1/3
- Active missions
  - ERS-2 (HBR, LBR)
  - ENVISAT (HBR,LBR)
- Future missions
  - GOCE
  - ADM
  - SMOS
  - CRYOSAT
  - Earthcare
  - SWARM
  - ...
  - SENTINELS
  - ...
- Non ESA missions facilities
- ESA missions facilities
  - UK-PAC
  - UK-PAF
  - I-PAC
  - I-PAF
  - D-PAC
  - D-PAF
  - E-PAC
  - CNES
  - IFREMER
  - LRAC
  - Esrin

✓Archive facility coordination issues:
- Data format, media, archive systems, migration plans ...

Some data archiving issues...

✓ **Volume of data** to be archived are exponentially increasing

✓ At present, ESA **mandate** is to maintain archives for 10 years only after mission end of life
- For a missions like ERS-1, this period is already over.

✓ ESA is proposing a coherent technical and managerial **strategy**
- through Ground-segment coordination body with CNES, ASI, DLR, EUMETSAT...
- coordination efforts started with some practical results

✓ The long-term **funding** strategy is to be found
- Space agencies use the funds from other on-going programme to maintain these historical datasets
- Such a scheme is no sustainable on the long-term

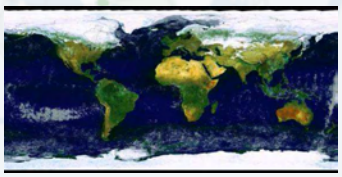
Some data preservation issues....

✓ **Data integrity** is not coherent across missions / facilities
- **Archive evolution** - data migration on routine basis, format definition, media maintenance, technology evolution, security, etc.
- **Completeness** - Integration of the physical telemetry payload data as acquired by at the stations with the related ephemeris (orbit...), auxiliary data (processing info to generate the nominal products...)
- **Logistic** - where the data are physically stored (processing may be performed in a separate location if required)
- **access** - information on what is archived and how to retrieve data (metadata, browse, products generation capability, products dissemination, local infrastructure, data reprocessing, etc.)

Some archive exploitation issues ....

✓ Provision of **reference consolidated data sets**...
- Systematic and periodical global dataset generation
- Assimilation / integration processes (fill the gaps)
- Science user support and acceptance
- Technical issues: e.g. standards
- On-demand access to simple and complex products

✓ **Evolution** of archive access
- Evolution of data processing algorithms
- ...



Actions taken for an European coordinated approach

- The ESA strategy is based on principles defined in
  - In agreement with ESA Delegation boards (DOSTAG, PB-EO
  - PB-EO Ground Segment Task Force recommendations
- Coordination started at ESA + National facilities via the G/S Coordination Group
  - Common archive standards (SAFE)
  - Define common data model (OAIS) and operational procedures
  - Standardization of interfaces
  - ...

Benefits for an European coordinated approach

- ✓ Cost reduction in terms of implementation and operations of existing and future projects
  - easier data exchange (based on same standards)
  - Better services to users
- ✓ Stronger European position at International level
  - Medium and Long term harmonized European data preservation policy
  - Leading voice in International Organizations, support coordinated approach also for GMES and GEOSS
- ✓ Optimal definition of infrastructure for common approach to data and knowledge preservation in Earth Science
  - Easier share of Facilities and related infrastructure
  - Ensure preservation of historical data, after the end of the mission lifetime

Thank you!

vincenzo.beruti@esa.int
luigi.fusco@esa.int